Using Machine Learning to Predict Student Success and Combat Inequity

Nathan Greenstein
Grant Crider-Phillips

April 6, 2022

UNIVERSITY OF OREGON | Academic Data Analytics

Welcome

Nathan Greenstein

Assistant Director of Machine Learning
Academic Data Analytics, University of Oregon

ngreenst@uoregon.edu

Grant Crider-Phillips

Machine Learning Analyst
Academic Data Analytics, University of Oregon

criderg@uoregon.edu

Academic Data Analytics

UNIVERSITY OF OREGON

Academic Data Analytics

- Office of the Provost
- Culture of data-driven decision-making
  - Shape policy
  - Prioritize equity
  - Increase transparency
- Focus areas:
  - Predicting student success
  - Understanding student feedback
  - Visualizing complex data
  - Understanding student and faculty progression

Session Roadmap

1. Project overview
2. Motivation
3. Our process
4. Early results and reflections
5. Discussion

Learning Goals

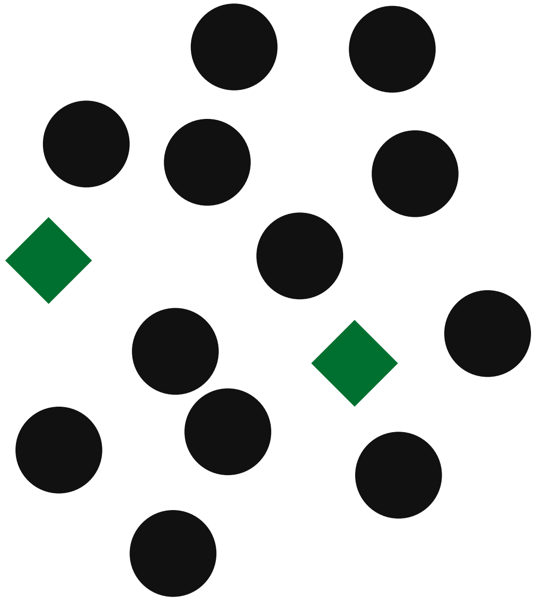
Understand applications of **machine learning**

Engage with interplay between machine learning and **equity**

Identify **implementation opportunities** at home institutions

Project Overview

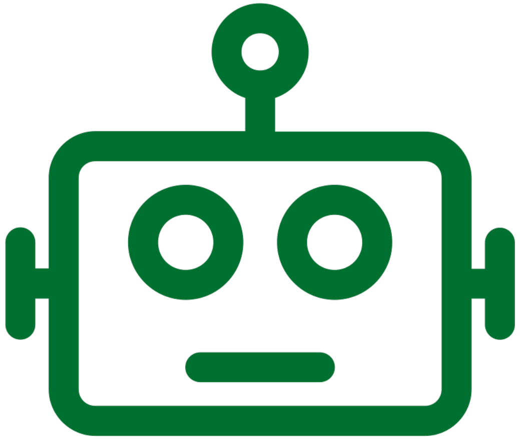
Prediction Task



Which incoming students will not persist to their second term?

- Include all incoming **first-time first-year students**
- Predict **before students matriculate**
- Each year, use predictions to **target early advising intervention**

- Many varieties; today's focus is **predictive analytics**
- Harnesses large amounts of **data** and **computing power**
- Searches for **relationships** between inputs and outputs
- Finds patterns **more complex** than human eyes and traditional methods can handle
- Not magical, but **powerful in the right situation**



Machine Learning

Motivation

●●○○○

Central Challenge

Non-Retention

- **Damaging** to students and university
- Disproportionately impacts **most vulnerable** students

Timely Intervention

- **Difficult to recover** from early negative experiences
- **Proactive interventions are more effective** than reactive ones

Finite Resources

- **Fewer advisors** than students
- **Must choose** who receives a given intervention first

Central Challenge

*Can we **predict which incoming students will not persist** to their second term?*

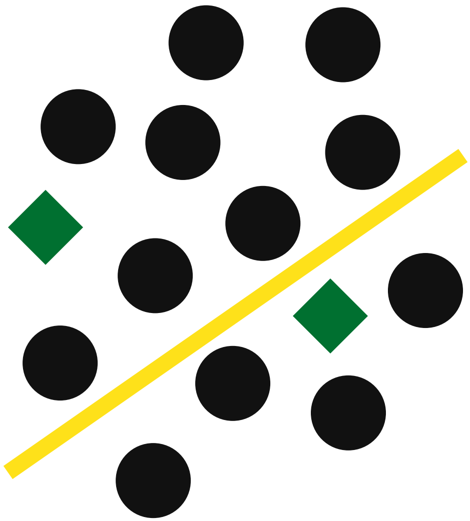
***How early** can we make our predictions?*

Non-Retention

Early Intervention

Finite Resources

- **Early advising** already in place
- **Mathematical model** already in use
  - Predicts **first-term GPA**
  - Traditional **linear regression**
  - **Unable** to predict second-term retention
- A useful tool, but a **compromise**
- Not evaluated for **equity**



Status Quo

Machine Learning

Promises

- Greater **predictive power**
- Better equipped for **challenging outcomes**
- Harnesses **bigger, messier data**

Concerns

- Will **human stakeholders** lose their voice?
- Might the algorithm be **biased or inequitable**?
- How much **transparency** will be offered?

Our Process

●●●○○

UNIVERSITY OF OREGON | Academic Data Analytics

Process Commitments

PARTICIPATORY

Engage meaningfully with a range of stakeholders

TRANSPARENT

Report honestly and accessibly on process and outcomes

EQUITY-ORIENTED

Apply lens throughout; demonstrably advance equity

Process Highlights

Participatory

- Partner closely with **Undergraduate Education and Student Success**
- **Converse** with other offices
- Reflect student body through **diverse data sources**

Transparent

- **Report actively** to UESS throughout
- **Publicly disseminate** methods and results
- Acknowledge **strengths and limitations**

Equity-Oriented

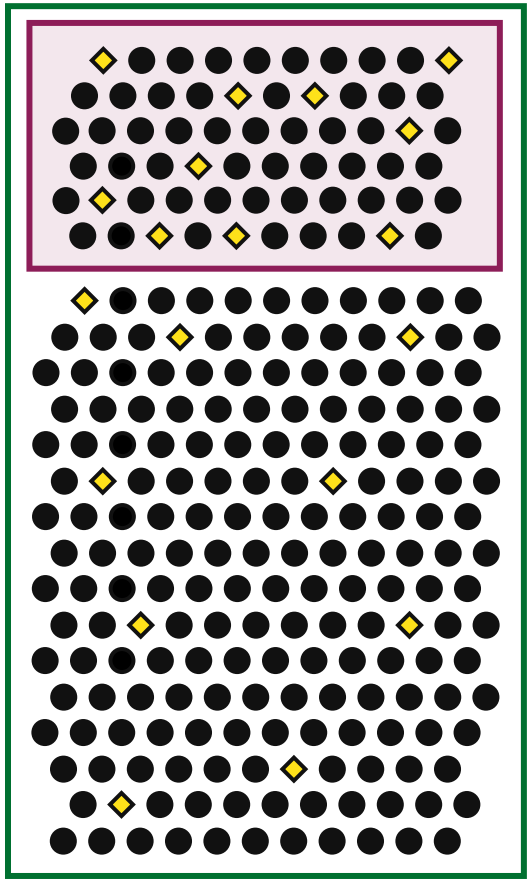
- With stakeholders, **define** equity standards
- Ground our work in **existing scholarship**
- Thoroughly **vet model** for equity and revise as necessary

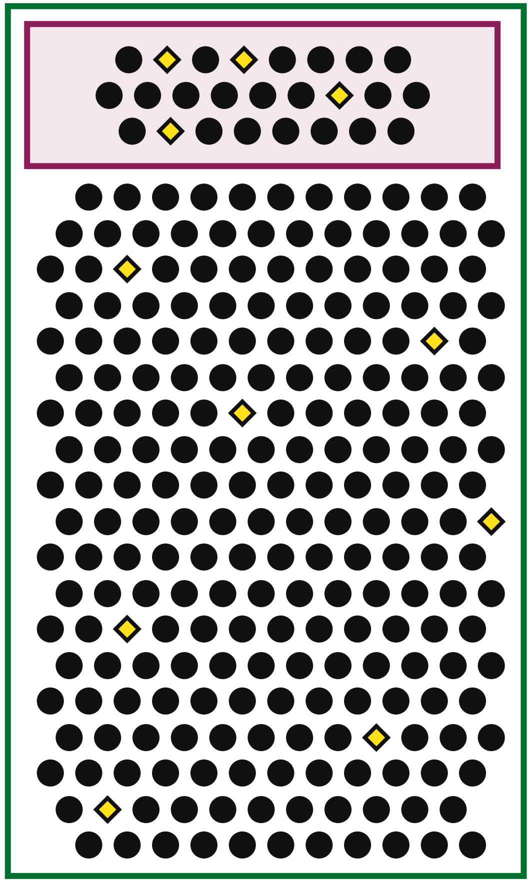
Early Results & Reflections

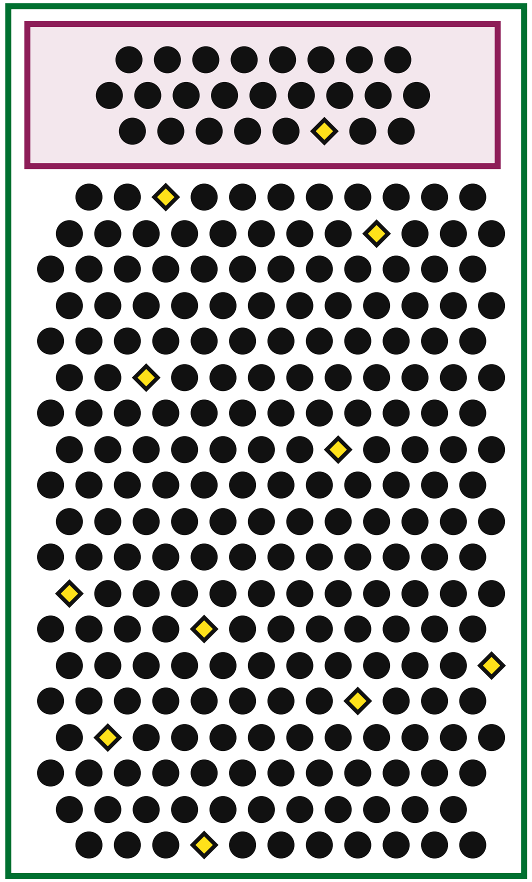
●●●●○

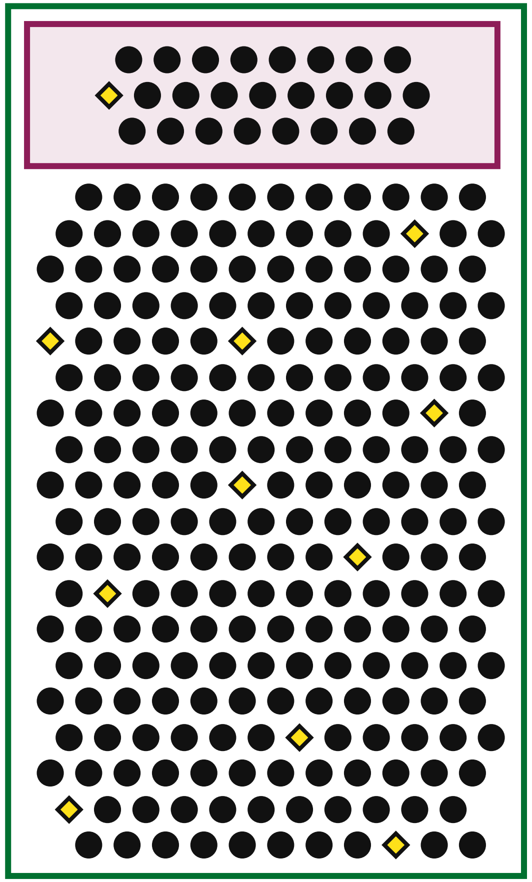
UNIVERSITY OF OREGON | Academic Data Analytics

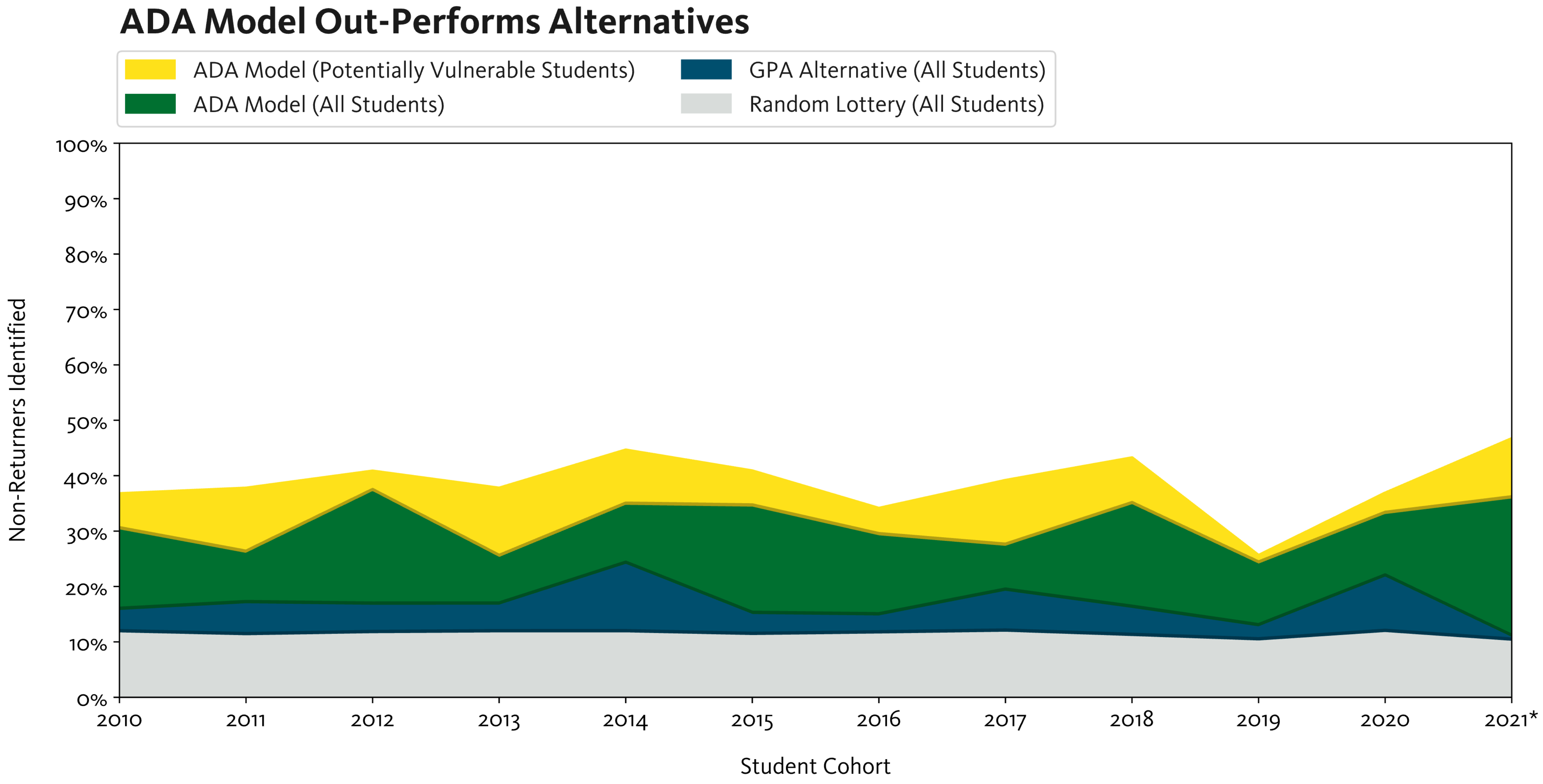
Model Performance, 2021 Cohort









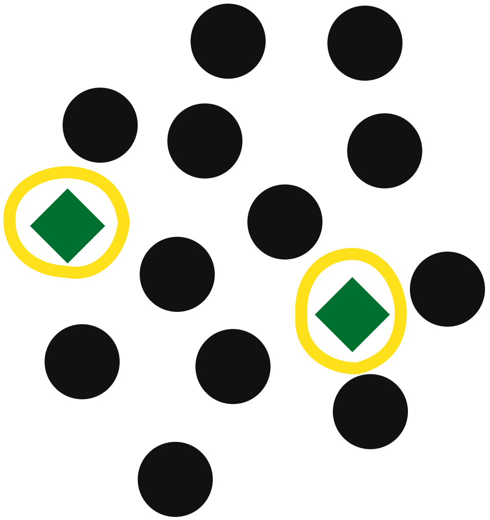


The 2021 cohort was hidden from the model during development to simulate an incoming cohort of students.

- Refine model **performance**
- Expand **equity analysis**; make any necessary **adjustments**
- **Deploy for incoming students** this year
- Begin **next predictive analytics project**

Ongoing Work

Reflections



- Confident that **performance exceeds alternatives**
- Room to **continue improving**
- Growing confidence in **model equity**
- **Process** was extremely successful
- Thoughtful approach, plus working in-house, enables **responsible machine learning**
- Ultimately, **harnessed powerful new tools without undermining** human stakeholders or potentially vulnerable students

Discussion

Session Roadmap

1. Project overview
2. Motivation
3. Our process
4. Early results and reflections
5. Discussion

Learning Goals

Understand applications of **machine learning**

Engage with interplay between machine learning and **equity**

Identify **implementation opportunities** at home institutions

Open Discussion

Thank you!

Nathan Greenstein
Grant Crider-Phillips

Academic Data Analytics
Office of the Provost
University of Oregon

https://provost.uoregon.edu/analytics

UNIVERSITY OF OREGON | Academic Data Analytics